Tracking colour object in presence of shadow

Hadj Hamma Tadjine

Institut für Informatik, Technische Universität Clausthal,
Julius-Albert-Strasse 4, 38678 Clausthal-Zellerfeld
tadjine@informatik.tu-clausthal.de

Segmentation is one of the most important steps leading to the analysis of images and image sequences. Its main goal is to divide an image into parts that have a strong correlation with objects or areas of the real world contained in the scene. A frequent occurrence in images is given by shadows. It is not difficult for human eyes to distinguish shadows from objects. However, identifying shadows by computer is a challenging research problem. Shadows are often integral parts of natural scenes and their identification is an important task in image and video analysis. On one hand, shadows are a valuable source of information about the shape and the relative position of objects in the scene. On the other hand, their presence may hinder the performance of object segmentation and interpretation systems for applications that require the identification of objects. Shadows can cause object merging, object shape distortion and even object losses (due to the shadow cast over another object). Moreover, the information about the shape and the colour of segmented objects may be distorted by shadows. Traditional motion-based tracking schemes cannot usually distinguish the shadow from the object itself, and this result in a falsely captured object shape. If we want to utilize the object's shape information for a pattern recognition task, this poses a severe difficulty. In this context arises the problem of identifying shadows and extracting useful information from the visual data they represent. The objective is to provide an efficient approach for the recognition of different types of shadows in both still images and video sequences. One of the main challenges in these applications is tracking the real object in a general video environment even if there is strong shadow influence (the shadow's effect is greatly attenuated). In video object tracking, shadows are commonly present for general lighting environments and present a confounding factor for correct object tracking. Tracking the real object in presence of shadow is a challenging problem. In this paper, we present a new method for objects tracking in a general video environment based on an illumination invariant colour space and a new external snake force called virtual electric field. Experimental results show that the proposed method can track the real object even if there is strong shadow influence.

1 Introduction

In ancient times, shadows were the antagonist of light. Light was present on one side of an object, and dark was present on the other side. After Alhazen (Abu Ali Al-

Hasan Ibn Al-Haytham), the view of light began to change, and with it the view of shadows.

The understanding of shadows in perceptual-psychological terms is only recently being explored. It is extremely important for object recognition that we are able to distinguish what is a shadow and what is not.

Shadow is an important problem, which receives relatively attention. In the case of computer vision, a shadows cause the erroneous segmentation of objects in the scene. It is not difficult for human eyes to distinguish shadows from objects. However, identifying shadows by computer is a challenging research problem [1, 2]. The recognition of shadow within a scene reveals a considerable amount of information about that scene. First, that there is at least one directional, localized light source in the scene; shadows are not present outdoors on over-cast days. Second, knowing the correspondence between a shadow and the object causing the shadow constrains the scene geometry [3].

Third, the difference in appearance between the same surface material lit and in shadow can tell one something about the difference between the direct light (light arriving along a direct line-of-sight from a light source, this is light that has not been reflected by surfaces in the environment.) and the light that illuminates the shadow.

Unfortunately, recognizing shadows in a scene is a difficult problem. Shadows can only be confidently recognized once the scene geometry, materials, and spectral flux (the light characterization at many points in the scene are known). This is more than just the characterization of sources of illumination because it includes the effects of inter-reflections between surfaces and the transmission properties of the environment. Knowing the scene's spectral flux and the material properties of a given surface, one can then deduce that a change in the appearance of the surface is due to a change in irradiance. With this knowledge and the determination that light from a source of illumination has actually been obstructed, one can conclude that a shadow is present. However, this is more knowledge than one can expect an observer to have when recognizing shadows in a scene. Detecting shadows falls into that large class of vision problems where, if most of the information about a scene is known, then the remaining information can be deduced from an image of the scene.

Shadows usually present a confounding factor for correct object tracking. Traditional motion detection schemes cannot distinguish the moving object and the shadows moving with it [2]. Therefore object tracking results based on traditional schemes usually result in a combination of the object and its shadow. This kind of result will pose severe difficulty if the contour is further passed to an analyser for object recognition. Eliminating the shadow and tracking the real contour of an object is a challenging problem.

A new application of snake model for general video object tracking is presented. A new external force is introduced into the snake equation based on the virtual electric field such that the active contour is attracted to a shape [4, 5]. The proposed method can deal with the problem of an object's ceasing movement temporarily, and can also avoid the problem of the snake tracking into the object interior.

Novel processing scheme is presented to project the image into an illumination invariant colour space such that the shadow's effect is greatly attenuated. The optical flow in this projected image together with the original image is used as a reference for object tracking so that we can extract the real object shape in the tracking process.

In this paper, the main aspect considered is the segmentation process. Therefore, the next section will present a brief overview of the existing active models, especially the snake models proposed by Kass et al. [6] and the gradient vector flow proposed by Xu et al. [7, 8]. Subsequently the main features of colour edge detection are discussed and finally some experimental results are presented.

2 Active contours

Image segmentation is an important component in image processing and computer vision. Deformable contour models using the snake approach have emerged as a powerful tool for contour extraction and edge detection [10].

The original active contour is related to Kass et al. [6], and is known as Snakes, due to the way the contour moves to its final position. In this model, the contour has an initial user-specified position and an associated objective function defined as the energy of the Snake which has to be minimised the energy. The energy is composed by different components ensuring some regularity properties (internal energy) and attracting it toward the a-priori interesting areas of the picture (external energy derived from image characteristics, such as the grey level gradient).

Let C(p): $[0,1] \to R^2$, $p \to C(p)$ be a parameterised planar curve, and let I: [0,a]x [0,b] $\to R^+$ be a given image in which we want to detect the objects boundaries. Associated to the curve C is the energy:

$$E = \int_0^1 \underbrace{\frac{1}{2}\left(\alpha\left|C'(p)\right|^2 + \beta\left|C''(p)\right|^2\right)}_{1} + \underbrace{E_{ext}\quad (C(p))}_{2} dp, \tag{1}$$

Where α and β are weighting parameters that control the snake's tension and rigidity, respectively. $C'(p)$ and $C''(p)$ denote the first and the second derivatives of C(p) with respect to p. The first part of the integral is related to the snake's internal energy and imposes restrictions to its movement by controlling α and β, respectively. The external energy function E_{ext} is responsible for driving the snake towards important features in the image. Thus, in a grey level image I(x,y), this external energy function can be written as:

$$E_{ext} = -\left|\nabla(G_\sigma(x,y) * I(x,y))\right|^2, \tag{2}$$

where $G_\sigma(x, y)$ is a two-dimensional Gaussian function with standard deviation σ, and ∇ is the gradient operator. This function is applied to the image in order to improve the image's edge map, and also to perform some noise reduction [8]. So, on regions closer to edges the gradient term yields higher values, and in consequence, stronger forces. A local minimum of equation (1) can be found by solving the Euler-Lagrange equation, i.e.

$$\alpha C''(p)-\beta C''''(p)-\nabla E_{ext}=0, \tag{3}$$

By treating C(p)=C(p,t) as a time dependent function, and assuming a solution is available if t=0, then the snake's evolving equation is given by:

$$\frac{\partial C(p,t)}{\partial t}=\alpha\, C''(p)-\beta\, C''''(p)-\nabla E_{ext}=0 \tag{4}$$

When the snake reaches a steady state, the term $\frac{\partial C(p,t)}{\partial t}$ disappears and a solution is found to equation (3).

Despite the model's consistency and simplicity, there are some performance problems associated to it [11, 15]. For example, the initial position of the snake must be close to the desired contour in the image, otherwise the snake may not evolve correctly, as it may find local minima away from the contour. Another common problem is that equation (4) can produced meaningless results is the snake happens to the cross itself. This problem can happen every time there are concavities or shape corners in the grey level image. Furthermore, the snake is indivisible, that is, it cannot split itself into two or more separated entities to adapt themselves to topological changes in the image domain.

2.1 Gradient vector flow

Most solutions define additional terms to be added to the external force component of equation (4), as a way to improve the snake's capture range and force it towards the image edges. Different external force models were proposed [13, 14]. This models also improves the initialisation problem, but does nothing to make the snake more adaptable to sharp corners and concave regions.

The latter problems could only be diminished by other external force model proposed by Xu and Prince [7], [9]. Their model, the Gradient Vector Flow is a bi-directional external force field that captures the object boundaries from either sides and can deal with concave regions. This model produces a field with strong forces near the edges, but also extends the gradient map farther into homogeneous regions using a computational diffusion process, which is also responsible for creating vectors that point into boundary concavities.

Gradient Vector Flow (GVF) method applied to snakes is a method that deals with concavity problem. To solve this problem, an external force must be acting on the snake and will pull it to concave object boundary. Suppose *f(x, y)* is the edge map of an image *I(x, y)*. Then the gradient of the edge map ∇f has vectors pointing towards the edges and these are normal to the edges on edge points. When this gradient force ∇f acts on the snake contour, it pulls the contour towards the concave boundary. However these gradient vectors have large magnitude only in the immediate vicinity of the boundary and are nearly zero at points away from the boundary. So the capture range of the snake will be very small. To increase the capture range the gradient map is extended to points away from the edges using a computational diffusion process. The gradient vector flow field $\upsilon(x,y) = (u(x,y), v(x,y))$ is derived from the following energy functional:

$$E(\upsilon) = \frac{1}{2}\iint (1 - g(|\nabla f|))((u - f_x)^2 + (v - f_y)^2)dxdy + \frac{1}{2}\iint g(|\nabla f|)(u_x^2 + u_y^2 + v_x^2 + v_y^2)dxdy \tag{5}$$

where $f(x,y)$ is defined as the edge-map function for a grey level image that has a higher value at the object boundary and μ is a blending parameter and g is a decreasing function of the gradient magnitude defined as follows:

$$g(x) = \exp(\frac{-x}{k}) \tag{6}$$

k is a positive constant controlling the smoothness of the resulting field. Calculus of variations is once again applied to Minimize (5) leading to the following Euler equations [8]:

$$g\nabla^2 u - (1 - g)(u - f_x) = 0 \tag{7}$$

$$g\nabla^2 v - (1 - g)(v - f_y) = 0 \tag{8}$$

where ∇^2 is the Laplacian operator. One solves (4) to obtain the GVF force field (u,v) that minimizes (3). More detailed investigation can be found in [8,9].

2.2 Virtual electric field

The presence of an electric charge produces a force on all other charges present. The charges exert a force on one another by means of disturbances that they generate in the space surrounding them. These disturbances are called electric fields. Each electrically charged object generates an electric field which permeates the space

around it, and exerts pushes or pulls whenever it comes in contact with other charged objects. For an electric charge Q placed at point $(0,0)$, the electric field $E(x,y)$ at point (x,y) is defined as follows:

$$E(x,y) = \left(\frac{Q.x}{4\pi\varepsilon r^3}, \frac{Q.y}{4\pi\varepsilon r^3} \right) \tag{9}$$

The electric field from multiple point charges can be obtained by taking the vector sum of the electric fields of the individual charges. If we have n charges, they make electric field E_n respectively and the total electric field E_{total} is given by the vector sum of each electric field as follow: $E_{total} = E_1 + E_2 + E_3 + ...+ E_n$. In traditional snake, the external force is defined as the negative gradient of external energy. In our work, this electric field is used as the external force F_{ext} instead of $F_{ext} = -\nabla E_{ext}$ in equation (3). The external force is directly defined as $F_{ext} = \frac{E_{total}}{|E_{total}|}$ without defining external energy. The electric field as the external forces is invoked by virtual charge Q. In image, the virtual charge Q is defined as image intensity itself or as the image edge at point (x,y) as follows [4]:

$$\begin{aligned} Q_{line}(x,y) &= -I(x,y), \\ Q_{edge}(x,y) &= -|\nabla I(x,y)|^2 \end{aligned} \tag{10}$$

As shown in Figure.1, the virtual electric field made by the virtual charges Q_{line} (around black pixels) which is used as external force has unlimited influence range and the same properties as the gradient vector flow.

If a point charge at (x,y) is $I(x,y)$, the electric field at (x_0, y_0) by point charge at (x,y) is as follows:

$$E(x,y;x_0,y_0) = E_x + E_y = \frac{c}{r^2}\left[\frac{x-x_0}{r}\hat{i} + \frac{y-y_0}{r}\hat{j} \right] \cdot I(x,y) \tag{11}$$

where $c = \frac{1}{4\pi\varepsilon}$ is a constant, $r = \sqrt{(x-x_0)^2 + (y-y_0)^2}$ the distance from the virtual charge and $I(x,y)$ is the intensity of image. At point (x_0, y_0) the x component E_x of the total electric force E_{total} by the electric charges in the region R is given as follows:

$$E_{total_x}(x_0,y_0;x,y)=\sum_{x,y\in R}E_x(x_0,y_0;x,y)=\sum_{x,y\in R}\frac{c}{r^3}(x-x_0)\cdot I(x-x_0)$$

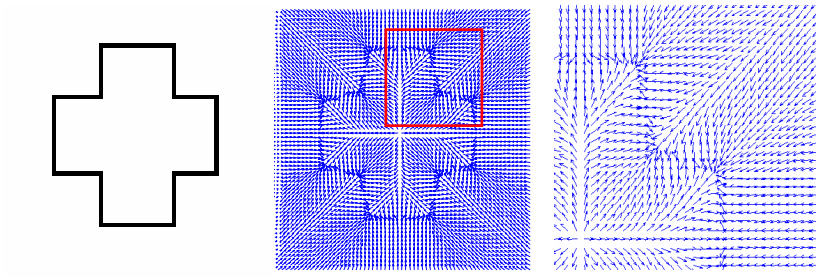
$$=\sum_{x,y\in R}\frac{-c(x_0-x)}{\left((x_0-x)^2+(y_0-y)^2\right)^{\frac{3}{2}}}\cdot I(x,y)$$

$$=h_x(x_0,y_0)*I(x_0,y_0). \tag{12}$$

It is clear that the total electric field E_{total} is calculated by the convolution of $I(x,y)$ and $h_x(x,y)$. By similarity, the y component E_y of the total electric field is given as follow (more detailed investigation can be found in [5]):

$$E_{total_y}=h_y(x_0,y_0)*I(x_0,y_0), \tag{13}$$

where $h_x(x,y)=\frac{-cx}{(x^2+y^2)^{\frac{3}{2}}}$ and $h_y(x,y)=\frac{-cy}{(x^2+y^2)^{\frac{3}{2}}}$.



(a) (b) (c)

Figure 1: Virtual electric field made by the virtual charges Q_{line}. (a) Image test, (b) virtual electric field $E/|E|$, (c) shown close-up within the boundary concavity.

3 Colour edge detection

In grey level imagery the extraction of edges is a two-step procedure. First the image is convolved with an operator usually based on the first or second derivative approximation of the Gaussian and then the response is threshold according to a criterion for assigning edge pixels. This general principle cannot be used directly for the colour imagery and instead the following methodologies have been used in the past [19]:

- The edges are detected in each component independently and their results are combined.
- The detection of edges is performed only on the luminance component and the result is integrated with heuristic rules based on the information coming from the rest of the components [20].
- The integrated gradient information's of all channels together is used for the detection of the edges [12], [16], [17].

The above methods have the following drawbacks:

- The first method suffers from the fact that is not clear how to combine the responses from the independently processed channels. It is clear that this processing does not exploit the correlation existing between the different components.
- For applying the second method, a priori information of external parameters such as surface reflectance characteristics and shadow should be available and hence is usually avoided [21].
- The third method presents advantages compared to the others since it integrates simultaneously colour information coming from the different components, This method is less sensitive to noise, and presents a solid theoretical background based on physical principles [18].

3.1 Colour gradient

A colour image (*R,G,B*) can be considered as a function (a vector field) which maps a two dimensional spatial information to a three dimensional (attributes) colour space. The gradient of this field can be generalised to the derivatives of the vector field. We follow the principled way to compute gradients in vector images as described by Silvano di Zenzo [22] and further used in [23], which is summarised as follows.

Let $\phi(x_1, x_2)$: $R^2 \to R^m$ be a m-band image with components for $\phi_i(x_1, x_2)$: $R^2 \to R$ for i =1,2,..,m (m=3 for colour images). Hence, at a given image location the image value is a vector in R^m. The difference at two nearby points $A = (x_1^0, x_2^0)$ and $B = (x_1^1, x_2^1)$ is given by $\Delta\phi = \phi(A) - \phi(B)$. Considering an infinite small displacement, the difference becomes the differential $d\phi = \sum_{i=1}^{2} \frac{\partial \phi}{\partial x_i} dx_i$ and its squared norm is given by:

$$d\phi^2 = \sum_{i=1}^{2}\sum_{j=1}^{2} \frac{\partial \phi}{\partial x_i}\frac{\partial \phi}{\partial x_j} dx_i dx_k = \sum_{i=1}^{2}\sum_{j=1}^{2} g_{ij} dx_i dx_j = \begin{bmatrix} dx_1 \\ dx_2 \end{bmatrix}^T \begin{bmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{bmatrix} \begin{bmatrix} dx_1 \\ dx_2 \end{bmatrix} \qquad (14)$$

Where $g_{ij} = \frac{\partial \phi}{\partial x_i}\frac{\partial \phi}{\partial x_j}$ and the extrema of the quadratic form are obtained in the direction of the eigenvectors of the matrix $[g_{ij}]$ and the values at these locations correspond with the eigenvalues given by:

$$\lambda_{\substack{max \\ min}} = \frac{g_{11} + g_{22} \pm \sqrt{(g_{11} - g_{22})^2 + 4g_{12}^2}}{2}, \tag{15}$$

with corresponding eigenvectors given by $(\cos\phi_{\substack{max \\ mix}}, \sin\phi_{\substack{max \\ mix}})$, where the angles $\phi_{\substack{max \\ mix}}$ are given (modulo π) by:

$$\phi_{\substack{max \\ min}} = \frac{1}{2}\arctan\frac{2g_{12}}{g_{11} - g_{22}}, \phi_{min} = \phi_{max} + \frac{\pi}{2} \tag{16}$$

Hence, the direction of the minimal and maximal changes at a given image location is expressed by the eigenvectors ϕ_{max} and ϕ_{min}, respectively. The corresponding magnitude is given by the eigenvalues λ_{max} and λ_{min}, respectively.

In comparison with grey level imagery there is a minimal rate of changes of the gradient perpendicular to the normal rate of change. In this case, the edge strength depends on both parameters. Therefore, the assignment of edge pixels in the colour domain depends on both rate changes [17], [18], [24].

The minimal rate of change could be different than zero. For example λ_{max} could be equal to λ_{min} which means that there is equal rate to all directions. Usually the relation between these parameters is given with a function expressing their dissimilarity (how λ_{max} compares to λ_{min}), for example by subtraction $\lambda_{max} - \lambda_{min}$ as proposed in [25], which will be used to define gradients in colour images.

3.2 Shadow

Shadow is an important problem, which receives relatively attention. In the case of computer vision, a shadows cause the erroneous segmentation of objects in the scene. It is not difficult for human eyes to distinguish shadows from objects. However, identifying shadows by computer is a challenging research problem.

Shadows occur when objects totally or partially occlude direct light from a source of illumination and can be divided into two classes: self-shadow and cast shadow [3]. A cast shadow is projected by the object in the direction of the light source; a self-shadow is the part of the object, which is not illuminated by direct light and is further classified in umbra and penumbra. The umbra corresponds to the area where the

direct light is totally blocked by the object, whereas in the penumbra area it is partially blocked [3].

The presence of cast shadows in an image can modify the perceived object shape, while the presence of self-shadows modifies the perceived object shape and its colour. In order to provide a correct description of the objects, shadows should be identified and classified. The classification into cast and self-shadows is based on the assumption that the intensity values of pixels in a self-shadow region are larger than those in the corresponding cast shadow region. This represents a limitation of the method since it leads to a misclassification if objects are significantly darker than the background or if a cast shadow receives light reflected from another object.

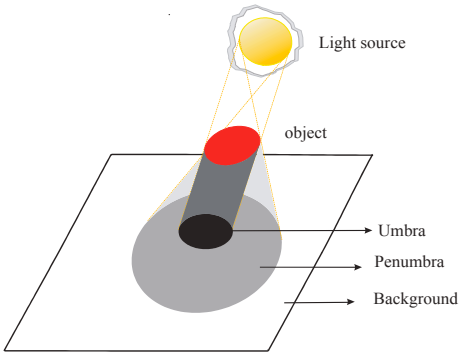


Figure 2: Nature of shadow.

Several methods for identifying shadows have been developed in recent years. Two different approaches have been followed, the first based on models which are used to represent the a priori knowledge of the three-dimensional geometry of the scene, the objects, and the illumination [26, 27], the second is more general and identifies shadows by exploiting their properties in geometry, brightness and colour [2, 28]. These approaches have thus a limited application range.

In this paper, we present an approach for colour object detection in presence of shadow, which overcomes some of the abovementioned limitations, namely the range of applicability with respect to the illumination conditions, and with respect to the need of an active observer. Since colour images provide generally more information than grey value images, more detailed edge information is expected from colour edge information [31]. The attention is directed towards the use of colour information because intensity-based edge detectors cannot distinguish between various transition types. The choice of colour features is of great importance for the purpose of proper image segmentation and induces equivalence classes to the actual segmentation algorithm.

3.3 Invariant coloured models

Colour information is exploited by considering colour features that show invariance properties with respect to changes in the illumination conditions that is to shadows and shading.

Let the colour gradient be denoted by ∇C, the proposed colour gradient for *RGB* (∇C_{RGB}) is as follows:

$$\nabla C_{RGB} = \sqrt{\lambda_{max}^{RGB} - \lambda_{min}^{RGB}} \tag{17}$$

for

$$\lambda_{\substack{max \\ min}} = \frac{g_{11}^{RGB} + g_{22}^{RGB} \pm \sqrt{\left(g_{11}^{RGB} - g_{22}^{RGB}\right)^2 + 4\left(g_{12}^{RGB}\right)^2}}{2}, \tag{18}$$

where:

$$\begin{aligned} g_{11}^{RGB} &= \left|\frac{\partial R}{\partial x}\right|^2 + \left|\frac{\partial G}{\partial x}\right|^2 + \left|\frac{\partial B}{\partial x}\right|^2, \\ g_{22}^{RGB} &= \left|\frac{\partial R}{\partial y}\right|^2 + \left|\frac{\partial G}{\partial y}\right|^2 + \left|\frac{\partial B}{\partial y}\right|^2, \\ g_{12}^{RGB} &= \frac{\partial R}{\partial x}\frac{\partial R}{\partial y} + \frac{\partial G}{\partial x}\frac{\partial G}{\partial y} + \frac{\partial B}{\partial x}\frac{\partial B}{\partial y} \end{aligned} \tag{19}$$

Similarity, the colour gradient ∇C_{rgb} is given by:

$$\nabla C_{rgb} = \sqrt{\lambda_{max}^{rgb} - \lambda_{min}^{rgb}} \tag{20}$$

Using the chromaticity coordinates in detection has the advantage of being more insensitive to small changes in illumination that are due to shadows. In addition to these, new invariant colour models $c_1c_2c_3$, $c_4c_5c_6$ and $l_1l_2l_3$ are proposed [30]. The $c_1c_2c_3$ invariant colour features are defined as follows:

$$c_1 = \arctan\left(\frac{R}{\max(G,B)}\right),\ c_2 = \arctan\left(\frac{G}{\max(R,B)}\right),\ c_3 = \arctan\left(\frac{B}{\max(R,G)}\right) \tag{21}$$

Furthermore, $c_4c_5c_6$ and $l_1l_2l_3$ are defined as follows [21]**:**

$$c_4 = \frac{R-G}{R+G},\ c_5 = \frac{R-B}{R+B},\ c_6 = \frac{G-B}{G+B}. \tag{22}$$

$$
\begin{aligned}
l_1 &= \frac{(R-G)^2}{(R-G)^2+(R-B)^2+(G-B)^2}, \\
l_2 &= \frac{(R-G)^2}{(R-G)^2+(R-B)^2+(G-B)^2}, \\
l_3 &= \frac{(G-B)^2}{(R-G)^2+(R-B)^2+(G-B)^2}.
\end{aligned} \tag{23}
$$

The colour gradient $\nabla C_{C_1C_2C_3}$, $\nabla C_{C_4C_5C_6}$ and $\nabla C_{l_1l_2l_3}$ respectively are given by: $\nabla C_{c_1c_2c_3} = \sqrt{\lambda_{\max}^{c_1c_2c_3} - \lambda_{\min}^{c_1c_2c_3}}$, $\nabla C_{c_4c_5c_6} = \sqrt{\lambda_{\max}^{c_4c_5c_6} - \lambda_{\min}^{c_4c_5c_6}}$ and $\nabla C_{l_1l_2l_3} = \sqrt{\lambda_{\max}^{l_1l_2l_3} - \lambda_{\min}^{l_1l_2l_3}}$. For our next calculations note that $\nabla I(x,y)$ in will be replaced by $\nabla \mathrm{C}$. Note that $c_4c_5c_6$, $l_1l_2l_3$ and $c_1c_2c_3$ varies with a change in material only, *rgb* with a change in material and highlights, and *RGB* with a change in material, highlights and geometry of an object [29, 30].

(a) (b) (c)

Figure 3: Colour edge detection. (a) Original image, (b) edge map of the luminance component, (c) colour edge map on invariant colour features containing only the object boundaries.

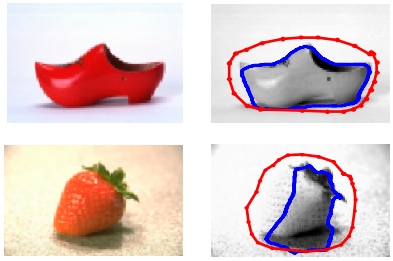
4 Experiments

Several implementation issues need to be addressed before evaluating the performance and robustness of the snake models, namely the test set, the initialisation conditions, the parameter values, the number of iterations, and the computation of the edge maps. The shape of the objects within these images contains smooth, high detail, convex and concave segments.

The parameter values are the same in all of our experiments, and they are equal to α=0.05, β=0, and λ=0.2, where α, β and λ are the elasticity, rigidity and regularization parameters of the GVF snake model, respectively. The number of iterations has been selected to be equal to 400.

4.1 Colour object detection

This section shows several examples of snake based on the virtual electric field force and gradient vector flow computations on real objects. The attention is directed towards the use of colour information because intensity based edge detectors cannot distinguish between various transition types [29].



(a) (b)

Figure 4: (a) The original image, (b) GVF snake, its initial position in red and the final position after 100 iterations in blue.

As intensity based edge detectors do not adequately distinguish between various transitions types, such as between different textures, reflections, objects and their shadows, etc., the use of colour information was investigated. The choice of colour features is of great importance for the snake. Colour features are selected enabling the snake algorithm to converge to the contours of object in images, which correspond to the material boundaries discounting the disturbing influences of illumination, shadows and highlights. The image analysis is performed in the (*R,G,B*) colour space. The size of the image is 128x128 (8 bits per colours).

A first observation, is that GVF snakes are limited in their application to colour images and do not result in a good detection of boundaries. The GVF snake segmentation results based on the intensity gradient is negatively affected by shadows and shading due to the irregular shape of the object. For these gradient fields it is not clear to which boundaries the snake contour should converge. Consequently, the final contour is biased and poorly defined [32]. The initial contour has partly converged to the wrong boundary (Figure 4).

The GVF snake segmentation results based on ∇C_{RGB} are poor due to disturbing influences resulting from unfavourable imaging conditions, such as shadows around the objects. In contrast, the final contours obtained by the GVF snake based on ∇C_{rgb}, $\nabla C_{C_4C_5C_6}$ and $\nabla C_{l_1l_2l_3}$ information, converge well to the true boundary. The results are shown in Figures 5 and 6.

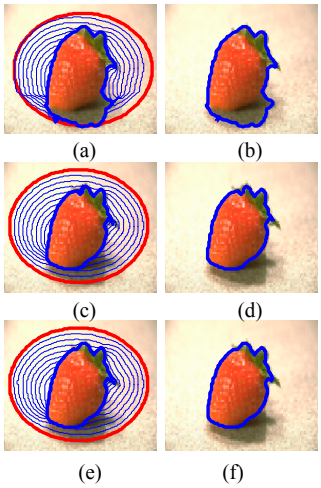


Figure 5: (a) Strawberry, (b) GVF snake based on ∇C_{RGB} its initial position in red, (c) and (d) based on $\nabla C_{I_1I_2I_3}$, (e) and (f) based on $\nabla C_{C_1C_2C_3}$.

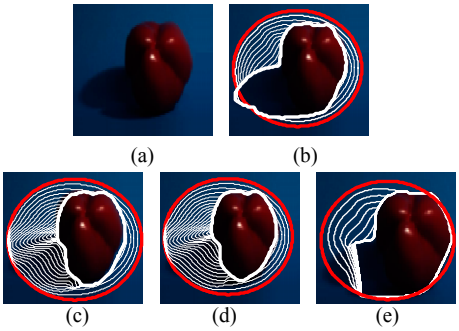
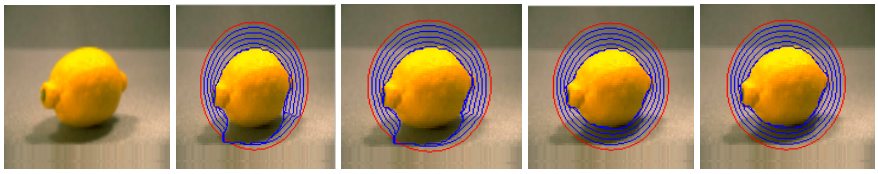


Figure 6: (a) Apple, (b) GVF snake based on ∇C_{RGB}, (c) based on ∇C_{rgb}, (d) based on $\nabla C_{C_1C_2C_3}$, (e) based on $\nabla C_{C_4C_5C_6}$.



(a) (b) (c) (d) (e)

Figure 7: Convergences of snake based on the virtual electric field force. (b) based on ∇C_{RGB} (c) based on $\nabla C_{C_4C_5C_6}$, (d) based on $\nabla C_{C_1C_2C_3}$, (e) based on $\nabla C_{l_1l_2l_3}$.

The second observation, which can be seen, is that The GVF as the external force has large captures range as well as convergence of concavities. However for calculation of the GVF, large computation time is required because the solution is computed by optimising diffusion equation. The proposed external force solved the problem above. For example, the typical calculation time of Gradient vector flow for the image in Fig.3 is 4,1 seconds using AMD Duron 850 MHz. On the other hand, the computation time using virtual electric field is 0.9 seconds on the same situation. Furthermore, on virtual electric field we can decide the capture range of the external force by limiting computational region of the image of reducing computation time [32].

The novelty of this work is to use a new external force for snake with colour gradient information in instead of using intensity gradient information steering the deformation process. However, using the proposed approach, better results are obtained and the external problem of disturbance is minimized (images are affected by shadows, Highlights).

4.2 Tracking

In visual tracking with snakes an accepted strategy for snake initialisation is as follows. Lead to capturing of colour object when the frame-to-frame this displacement is after capturing the object in the current frame, take the position of this snake and place it on the next frame [11]. Following such strategy with a GVF snake does not high or equivalently the temporal resolution is low, since the standard GVF snake fails if the initial snake does not include the object medial axis [33].

This deficiency leaves a scope to enhance the original GVF force towards making it robust to snake initialisation. Thus a snake-force *(u,v),* constraining the partial differential equation based generalized gradient vector flow through Dirichlet type boundary condition's is proposed as follows:

$$\begin{aligned} when \qquad (x,y) \in (A-B) : & \; g\nabla^2 u - (1-g)(u - f_x) \\ & \; g\nabla^2 v - (1-g)(v - f_y) \end{aligned} \tag{24}$$

$$when \qquad (x,y) \in \partial A : \nu = (u(x,y), v(x,y)) \tag{25}$$

$$when \qquad (x,y) \in \partial B : \nabla \nu = 0. \tag{26}$$

where ν is the direction of the object's movement, A is the image domain with boundary ∂A and B is the region bounded by the initial closed snake with boundary ∂B.

It is to be noted that the Boundary condition for the initial snake boundary ∂B is of Dirichlet type while that for the image boundary ∂A is of Neumann type [33]. However, the only constraining boundary condition is the Dirichlet one and which is responsible for the proposed enhancement. Adding such a boundary condition improves tracking the colour object with the presence of shadow. Several examples of GVF field computations on real objects are given. Using the proposed approach, improved results are obtained and the influence of image disturbances, such as shadows, is minimized (cf. Figures 8 and 9).

Figure 8: Ball tracking without shadow.

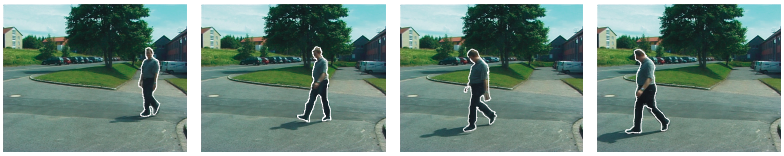


Figure 9: Outdoors body tracking.

5 Conclusions

In this paper results are presented, which show that the use of GVF snakes for detecting object edges in video can offer distinct advantages. A new external snake's force called virtual electric field is introduced. This force based on the electric field that is made by virtual electric charges on the boundary of an object in image. This paper shows how by using the proposed colour snake it is possible to take advantage of the little evidence in order to have a better detection of object. The performance of

the proposed snake based on colour gradient shows that the intensity-based snake is dramatically outperformed by the presented colour snake.

Also the time for computing the virtual electric field is much less than the time for the gradient vector flow. This shorter processing time is due to the computing of the virtual electric field, which can be accomplished by the simple calculations as convolution. Because the calculation of convolution is supported by modern computing technology as DSP (Digital Signal Processing) the calculations of the virtual electric field can be more efficient.

An efficient algorithm for tracking objects with shadows is presented. The algorithm can eliminate the distracting influence from the shadows and track the shape of the real object. The quality of the results can be significantly improved by employing the colour gradient. This can be very useful for higher-level vision processing. Future research for this project includes speeding up the algorithm such that it can work in real time applications.

6 Acknowledgement

The images used are available from Macmillan Publishing and from ISIS Project (University of Amsterdam). The authors express their grateful thanks to the *DAAD* (Deutscher Akademischer Austauschdienst) for financial support of this work.

References

[1] R. Cucchiara, C. Grana, M.Piccardi, and A. Patri.: Detecting objects, shadows and ghosts in video steams by exploiting colour and motion information, *Computer Vision and Pattern Recognition conference*, 2001, Kauai, Hawaii, USA, vol.2, pp.571-576.

[2] A. Prati, I. Mikic, C. Grana, M.M. Trivedi.: Shadow detection algorithms for traffic flow analysis: a comparative study, *in Proceedings of IEEE Intelligent Transportation System Conference*, 2001,Oakland, CA, USA, pp. 340-345.

[3] G. Funka-Lea, R. Bajcsy.: Combining Color and Geometry for the Active, Visual Recognition of Shadows, *Proc. International Conference on Computer Vision,* 1995, 203–209.

[4] H.K. Park, and M.J. Chung.: A new external for active contour model: Virtual electric field, *International Conference on Visualization, Imaging, and Image processing*, September, 2002, Malaga, Spain, pp. 85-89.

[5] H.H. Tadjine, G. Joubert, and S.M. Mouattamid.: Colour object detection with shadow using virtual electric field, *Proceedings of the International conference on image and signal processing (ICISP 2003)*, 25-27 June, 2003, Agadir, Morocco, vol. 1, pp. 53-60.

[6] M. Kass, A. Witkin, and D. Terzopoulos.: Snakes: Active Contour Models,

International Journal of Computer Vision, *1(4),* 1998, 321-331.

[7] C. Xu and J. L. Prince.: Gradient vector flow deformable models, in I. Bankman (Ed) *Handbook of Medical Imaging* (Academic Press, 2000), 159-169.

[8] C. Xu and J.L. Prince, Gradient vector flow: A new external force for snakes, *Proceeding of the IEEE Computer Vision and Pattern Recognition*, 1997, 66-71.

[9] C. Xu and J.L. Prince.: Snakes, shapes, and gradient vector flow, *IEEE Transactions on Image Processing*, 7(3), 1998, 359-369.

[10] M.D. Fairchild.: *Color Appearance Models* (Addison-Wesley, 1998).

[11] A. Blake, and M. Isard.: *Active contours: The application of techniques from graphics, vision, control theory and statistics to visual tracking of shapes in motion,* (Springer, 1998).

[12] T.N. Pappas, An adaptive clustering algorithm for image segmentation, *IEEE Transactions on Signal Processing*, 40, 1992, 901-914.

[13] L.D. Cohen, on active contour models and balloons, *Computer Vision Graphics and Image Processing: Image Understanding*, 53(2), 1991, 211-218.

[14] L.D. Cohen, and R. Kimmel.: Global minimum for active contour models: A minimal path approach, *International Journal of Computer*, *24(1),* 1997, 57-78.

[15] G. Sapiro, R. Kimmel, and V. Caselles.: Geodesic active contours, *International Journal of Computer Vision*, *22(1)*, 1997, 61-79.

[16] G. Sapiro, and D.L. Ringach.: Anisotropic diffusion of multivalued images with applications to color filtering, *IEEE Transactions on Image Processing*, 5(11), 1996, 1582-1587.

[17] G. Sapiro.: Vector valued active contours, *IEEE Conference on computer Vision and Pattern recognition*, 1996.

[18] G. Sapiro.: Color snakes, *Computer Vision and Image Understanding*, 68(2), 1997, 247-253.

[19] E. Saber, A.M Tekalp, and G. Bozdagi.: Fusion of color and edge information for improved segmentation and edge linking*, Image and Vision Computing*, Vol.15, N_o.10, pp 769-780, 1997.

[20] S. J. Sangwine, and R.E.N. Nornee.: The colour image-processing Handbook (*Chapman & Hall*, 1998).

[21] M.D. Levine.: Vision in Man and Machine (*McGraw -Hill*, 1985).

[22] S. di Zenzo.: A note on the gradient of multi-images, *Computer Vision Graphics and Image Processing*, 33, 1986, 116-125.

[23] H.C. Lee, and D.R. Cok.: Detecting boundaries in a vector field, *IEEE Transaction on signal processing*, 39(5), 1991, 1181-1194.

[24] P. Bloomgren, and T. Chan: Colour TV, Total Variation methods for Restoration of vector Valued Images, *IEEE Transactions on Image Processing*, 7(3), 1998, 304-309.

[25] D. Cummani, and P. Meer.: Edge detection in multispectral images, *Computer Vision, Graphics and Image Processing: Graphical Models Image Processing*, 53(1), 1991, 40-51.

[26] M. Bejanin, A. Huertas, G. Medioni and R. Nevatia.: Model Validation for

Change Detection. *2nd International IEEE Workshop on Applications of Computer Vision*, 1994, Los Alamitos, USA, pp.160-167.

[27] J. -P. Thirion.: Realistic 3D Simulation of Shapes and Shadows for Image Processing, *CVGIP: Graphical Models and Image Processing*, 1992,54 (1): pp. 82-90.

[28] C. Jiang and M. O. Ward.: Shadow Segmentation and Classification in a Constrained Environment, *VGIP: Image Understanding*, 1994, 59 (2), pp. 213–225.

[29] T. Gevers, and A.W.M. Smeulders.: Colour based Object recognition, Pattern Recognition, 32(3), 1999, 453-464.

[30] T. Gevers, and A.W.M. Smeulders.: PickToSeek: Combining colour and shape invariant features for image retrieval, *IEEE Transactions on Image Processing*, 9(1), 2000, 102-119.

[31] E. Salvador and T. Ebrahimi.: Cast shadow recognition in color images. In Proc. 11th European Signal Processing Conference, September 2002, Toulouse, France.

[32] H.H. Tadjine, and G. Joubert.: Novel framework for gradient vector flow, *Proceedings of the second IASTED International Conference on Visualization, Imaging, and Image processing*, September, 2002, Malaga, Spain, pp. 52-57.

[33] C.A. Hall, and T.A. Porsching.: Numerical analysis of partial differential equations, *Englewood Cliffs, New jersey: Prentice Hall*, 1990.